AD-A074 936 PENNSYLVANIA STATE UNIV UNIVERSITY PARK APPLIED RESE--ETC F/G 5/10
AN INVESTIGATION OF DUAL SENSORY STIMULUS PRESENTATION OF COMPL--ETC(U)
JUN 79 A BARBOUR N00024-79-C-6043
UNCLASSIFIED ARL/PSU/TM-79-115 NL

1 OF 2
ADA
074936

SIFIED

1 OF 2

ADA
074936

MICROCOPY RESOLUTION TEST CHART

AD A074936

LEVEL II

AN INVESTIGATION OF DUAL SENSORY STIMULUS PRESENTATION OF COMPLEX NOISE-LIKE SOUNDS

Alfred Barbour

Technical Memorandum
File No. TM 79-115
June 9, 1979
Contract No. N00024-79-C-6043

Copy No. 5

The Pennsylvania State University
Institute for Science and Engineering
APPLIED RESEARCH LABORATORY
Post Office Box 30
State College, PA 16801

NAVY DEPARTMENT

NAVAL SEA SYSTEMS COMMAND

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

14

1. REPORT NUMBER: ARL/PSU TM-79-115

2. GOVT ACCESSION NO.: 9 Master's thesis

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle): 6 AN INVESTIGATION OF DUAL SENSORY STIMULUS PRESENTATION OF COMPLEX NOISE-LIKE SOUNDS,

5. TYPE OF REPORT & PERIOD COVERED: MS thesis, August 1979

6. PERFORMING ORG. REPORT NUMBER: TM 79-115

7. AUTHOR(s): 10 Alfred/Barbour

8. CONTRACT OR GRANT NUMBER(s): 15 N00024-79-C-6043

9. PERFORMING ORGANIZATION NAME AND ADDRESS:
The Pennsylvania State University
Applied Research Laboratory
P. O. Box 30, State College, PA 16801

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME AND ADDRESS:
Naval Sea Systems Command
Department of the Navy
Washington, DC 20362

12. REPORT DATE: 11 9 June [redacted]79

13. NUMBER OF PAGES: ~~95~~ pages & figures 12 96

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)

15. SECURITY CLASS. (of this report): Unclassified, Unlimited

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release, distribution unlimited, per NSSC (Naval Sea Systems Command) 6/26/79

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The ability of a listener to detect changes in auditory and visual signals was investigated. Subjects were presented auditory signals via headphones, a spectral representation of the signal via a CRT, or both representations simultaneously. The signal pairs were bands of noise buried in a noise. The Signal-to-Noise (SNR) increased slowly. This gave the signal the appearance of emerging out of the noise. The task of the subjects consisted of determining which of two possible signals were presented in the noise and

DD FORM 1473, 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. ABSTRACT (Continued)

responding by pressing a designated response button. Subjects were asked to use the criterion of reasonable certainty. The factors of interest were the mean SNR at which the subject was able to make a discrimination and to respond and the probability of a correct response. The mean SNR's for each of the modes of presentation were compared to determine the significance of combining sensory modalities in signal detection and discrimination tasks.

Five university students were used as subjects in the study. The results indicate that for two of three signal patterns used in this experiment, the combined audio-visual presentation mode is superior to either the auditory or the visual modes used singly, while for one of the signal pairs used, an amplitude modulated signal, the auditory presentation mode yielded the best performance.

Abstract

The ability of a listener to detect changes in auditory and visual signals was investigated. Subjects were presented auditory signals via headphones, a spectral representation of the signal via a CRT, or both representations simultaneously. The signal pairs were bands of noise buried in a noise. The Signal-to-Noise Ratio (SNR) increased slowly. This gave the signal the appearance of emerging out of the noise. The task of the subjects consisted of determining which of two possible signals was presented in the noise and responding by pressing a designated response button. Subjects were asked to use the criterion of reasonable certainty. The factors of interest were the mean SNR at which the subject was able to make a discrimination and to respond and the probability of a correct response. The mean SNR's for each of the modes of presentation were compared to determine the significance of combining sensory modalities in signal detection and discrimination tasks.

Five university students were used as subjects in the study. The results indicate that for two of three signal patterns used in this experiment, the combined audio-visual presentation mode is superior to either the auditory or the visual modes used singly, while for one of the signal pairs used, an amplitude modulated signal, the auditory presentation mode yielded the best performance.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

The author gratefully acknowledges the contributions of Dr. C. P. Janota throughout this work. The major portions of the test apparatus were designed by Dr. Janota. Also, his technical assistance, suggestions regarding the choice of stimuli, and the analysis of the results were paramount to the completion of this thesis.

Also, the author expresses his appreciation for the suggestions and encouragement of Dr. L. G. Kerr, his thesis advisor. In addition, the assistance of fellow graduate students D. W. Martin, who is responsible for much of the computer software used in the data reduction procedures, and P. T. Cornell, who aided in the data analysis, is greatly appreciated.

Finally, the assistance furnished by the Applied Research Laboratory of The Pennsylvania State University under contract with the Naval Systems Command is greatly appreciated. The work was sponsored by the Naval Air Systems Command (AIR 370).

CHAPTER I

INTRODUCTION

1.1 General

Man gains the majority of information about his environment through his two primary senses, those of vision and audition. In our modern-day society it is rare to find a source of input information to one sense modality that is completely isolated from the other. The interplay of these two senses are used in our everyday lives, for both occupation and recreation. A machinist uses these senses to monitor the functions being performed by his machines. Likewise, these senses are used to enjoy a host of recreational activities from watching a football game to enjoying an opera.

The interaction of these two sensory modalities is a topic that has interested psychologists and other researchers for many years. Many studies have addressed the topic of sensory interaction in regard to human information processing in vigilance and detection performance studies. Since the eye and ear function as independent informational channels, each is capable of receiving information concurrently with the other. The question of sensory interaction arises, that is, whether the information in the component modalities summate in an informational sense or whether the process is a modification of the component information either by intersensory masking or by a facilatory effect, and remains unanswered.

The initial findings of Schafer and Shewmaker in 1953 that showed the general superiority of combined audio-visual presentation, have been supported in more recent studies. Loveless (1957) compared single modality and combined audio-visual signals in short watch sessions and found a higher detection rate for the combined signals. Bruckner and Mc Grath (1961) reported an increased detection probability for redundant signal presentation to both the auditory and visual senses.

1.2 Statement of the Problem

The experiment whose results will be presented in this thesis will, it is hoped, add to the knowledge of how a human detects and discriminates between complex nonspeech sounds presented both aurally and in a visual representation. The results of the psychoacoustic experiment will be reported and analyzed in terms of information-processing theory and signal detection theory. Models of feature extraction and stimulus perception will be contrasted for each sense modality, and an interaction theory will be discussed for the combination of these two modalities. The models assume a similarity in higher order processing of the two modalities, since both the eye and ear function as transducers to convert one form of energy to another (Stevens, 1958).

The eye and ear operate separately as independent informational channels, each capable of receiving energy adapted to their function. Since the human is capable of receiving simultaneous inputs to different sense modalities, the two major senses, sight and sound, have been studied at great length by researchers in the past.

An environment where audition is of major importance is that of the sonar operator. The sonar operator relies heavily on the auditory sense to detect and discriminate among ocean-borne sounds. Many researchers have studied the sonar environment and proposed modeling strategies. These models are not intended to be descriptions of the way human observers operate, but merely as normative models against which their performance may be compared (Janota, 1977). Previous researchers have examined the role of the human listener in detecting and discriminating ocean-borne sounds. Ocean-borne marine sounds are generally broadband and may exhibit some significant tonal qualities. Janota (1977) investigated the detectabilities of a number of laboratory-generated and actual recorded ocean sounds and tested the detectability and discrimination of these. Martin (1978) also studied the detectability of broadband signals with specific interacting features.

Many researchers have elected to use dual presentation to both the visual and auditory sense modalities to try to enhance the detection of acoustic signals. Since auditory signals do not lend themselves to visual processing, an energy transform must be performed. Following this energy transform, a display system must be *invented* that will give an interpretable realistic visual representation of the acoustic energy. The majority of the displays used previously seem unconventional and unnatural, that is, uncomfortable to use compared to displays that are typically used. By convention, engineers, technicians, and acousticians have viewed acoustic signals in the form of an electrical representation on an oscilloscope. For that reason, an oscilloscope was used for the visual display in this experiment.

1.3 Approach

The experiment reported in this thesis was designed to investigate the efficacy of using redundant information presented to both the auditory and visual sense modalities in an effort to aid in the detection and discrimination of signals. The experiment involved three different sound pairs which subjects were asked to detect and discriminate between. The signals were presented aurally, visually, and with both modes combined. The detection task involved detection of the signal against a background of white noise. The discrimination task consisted of determining which of a previously learned pair of signals was presented against the noise. The auditory signals used were laboratory-generated sounds containing various fixed and dichotomous features. A dichotomous feature is defined as a characteristic that is present in one member of the pair, but absent in the other. These features included stationary octave bands of noise centered at several frequencies and the presence or absence of amplitude modulation in the noise bands. The visual signals were a direct transform of the acoustic energy into electrical energy and displayed on an oscilloscope.

The experimental procedure used has been termed the *Modified Threshold Technique* (Janota, 1977) and involves a sequential classification task in which the signal-to-noise ratio (SNR) of the stimulus increases with time on each given trial. The subject responds during the trial when he is confident as to which signal is being presented. That is, if the subject does not feel reasonably confident with the signal choice to make a decision, he can wait for additional information about the features of the signal, since the SNR is increasing, before committing himself. The data of interest are: 1) the signal-to-noise ratio at which the subject is willing to respond and 2) the percentage of correct responses for a given signal pattern. By knowing these parameters, the discrimination performance for specific signal patterns can be determined and the relevant signal features required for recognition can be assessed.

The experiment was conducted in three parts or modes. In mode 1, subjects received auditory information only; in mode 2, subjects received visual information only; while in mode 3, the subjects received both the auditory and visual information simultaneously. Prior to detailing the experiment to be reported, a brief review of the literature relevant to this topic is presented in Chapter II. The studies cited from the literature used dual modality presentation of redundant information. The techniques used, the types of signals, and the visual displays employed will be discussed.

Chapter III will present various information processing models of bisensory information presentation. The models cited have all used auditory and visual information to assess the utility of presenting information to more than one modality. A proposed model of the discrimination process will also be discussed along with the hypothesis this experiment was designed to test.

Chapter IV will discuss the experimental paradigm used to test the hypothesis previously stated. A description of the noise-like sounds will be given and the reasons for the choice of these sounds. The choice of signal pairs used and the type of visual display will also be noted. The methods for selection and training of subjects will be discussed, as will the *modified threshold technique.* The methods of data reduction will also be outlined, with a discussion of the important parameters, SNR at response, and the probability of a correct response *P(C).*

Chapter V will present the results obtained from the experiments. The results are contrasted against the information processing models discussed in Chapter III. Experimental factors which may have led to subject bias affecting the decision criteria will be discussed, and methods for accounting for and handling these biases will be detailed.

Chapter VI will present a summary of the major findings of this study and detail some of the conclusions which may be drawn from these findings.

CHAPTER II

REVIEW OF INVESTIGATIONS OF BISENSORY PRESENTATION OF INFORMATION

2.1 General

An area of psychological research which has received much attention is the assessment of the advantage to be gained by simultaneous use of more than one sensory modality as information channels. There are two situations when these techniques may aid the human operator in assimilating information. The first of these is the case where separate message streams of information are delivered to each modality with the intent of increasing the amount of information handled per unit time. The second is where the incoming stimuli are difficult to detect, recognize, or discriminate from irrelevant or masking stimuli. It is this second situation that is the main concern of this thesis.

The presentation of partially or totally redundant information to more than one sense will, to a point, aid the detection of weak stimuli. That is, the combined condition will be better than the better of the two single modalities, but less than the arithmetic sum.

2.2 Synopsis of Previous Research

Bruckner and Mc Grath (1961) compared detection performance on vigilance tasks designed to test the auditory and visual modes both in single and dual mode presentation tasks. The study involved using different degrees of redundancy in the presentation of information. The visual task of the subjects was to detect an increment in brightness of a continuous light source viewed through a frosted glass. The auditory task required subjects to detect increments in loudness of a 750 Hz continuous tone presented via headphones.

The subjects were tested in a variety of conditions. The subjects were tested on the visual task only, the auditory task only, and both modes combined. During the visual-only task, subjects monitored only the visual display. For the auditory-only task, subjects monitored only the auditory display. In the dual mode presentation case, subjects were required to monitor both display systems with signals appearing simultaneously on each. This was termed the redundant task. The next condition required subjects to monitor both display systems, but with signals appearing on either but not both. This condition was termed the non-redundant task. The final condition required the subjects to monitor both displays simultaneously. One-third of the signals appeared on either display but not both. One-third of the signals appeared on both displays simultaneously; this task was termed the partially redundant case.

The subjects were presented signals at the rate of six per quarter-hour period; however, the interstimulus interval varied from 9 seconds to 5 minutes. The results showed that in terms of percentage of detections, the redundant task was superior to the other conditions tested.

Osborn, Sheldon, and Baker (1963) required subjects to detect interruptions in a continuous light source, a continuous sound, or both sources during a 3 hour watch session. They reported the mean detection rate at 30 minute intervals. The results showed the redundant case to be superior to either the auditory or the visual modalities for the detection of the signals used.

Corcoran and Weening (1969) tested subjects on the detection of signals by requiring subjects to detect four signal patterns presented in the auditory mode, visually, or with both modes occurring simultaneously. The signals used were narrow-band signals consisting of two frequencies of 1001 and 1201 Hz with two different beat rates of 2 or 3 beats per second. These signals were presented to the subjects against a thermal noise background. The signals were presented over headphones in the auditory case, via an oscilloscope in the visual case, and in both conditions simultaneously in the audio-visual case. Again, the findings show that the redundant presentation case was superior in detecting the signals.

Halpern (1970) and Wells (1971) used various recorded broadband marine sounds for auditory stimuli. These sounds were presented to subjects via headphones in the auditory presentation case. In both of these experiments, the visual display consisted of a 4 X 4 matrix of circles displayed on a screen. Each of the 16 circles was attached to the output of a bandpass filter tuned to a specific frequency in the auditory input signal range. The circles were animated by the output of the filters, that is, the excitation of a filter would cause its corresponding circle to increase and decrease in size. The circles were arranged such that the higher frequencies excited the upper left-hand circles while the lower frequencies excited the lower right-hand circles. From the fluctuations in the size of the circles, the subject could gain information as to the spectral shape of the input signal. It was found by Halpern (1970) and Wells (1971) that detection of the broadband signals was better in the dual presentation case than in either presentation mode used singly.

A study by Colquhoun (1975) used four amplitude modulated tones of 300, 500, 700, and 900 Hz for the auditory stimuli and four concentric brightness-modulated rings, one corresponding to each of the frequencies as the visual display. The rings were displayed on a short persistence Cathode Ray Tube (CRT) oscilloscope. The subjects were required to detect the signals visually, aurally, and with both modes combined. Colquhoun concludes that, where efficiency both in the initial detection of targets and their subsequent identification and tracking are equally important, the best solutions would seem to be to retain both auditory and visual displays and to ensure that these are monitored concurrently.

Although in all the experiments cited above, the dual mode detection performance has been found to be superior to that of single mode, a caution must be drawn. The added benefit in detection is not equivalent to the combination of the detection rates from each modality. Bruckner and Mc Grath (1961) and Loveless (1970) have found that combining the detectabilities of the single modalities overestimates the dual mode prediction. This point will be discussed further in the following chapter.

CHAPTER III

BISENSORY PRESENTATION, AN INFORMATION-PROCESSING APPROACH

3.1 General

The question of whether information presented to more than one sense modality simultaneously can be combined within the nervous system or cognitive processes to yield a greater efficiency or level of performance over single mode presentation has been investigated in the past. Of particular interest and relevance to this thesis are the studies that have investigated the combined use of the auditory and visual senses either in vigilance or detection tasks. More specifically of interest are the studies which used bimodal presentation of nonverbal stimuli. The question of the summation of information arises. It has been shown by a number of studies that the detection of weak or masked signals can be enhanced by presentation to more than one modality. Many models have been proposed to explain these findings.

3.2 Proposed Models

Loveless, Brebner, and Hamilton (1970) propose a *statistical summation* model which suggests that, since the eye and ear are independent channnels, each channel arrives at a detection decision independently. These independent decisions are then passed on to a second decision-making stage where the final decision is made. It is proposed that if either or both channels conclude that a signal is present, then the observer reports that a signal was indeed present. Loveless et al. (1970) propose the following equations as a model of the process:

$$P_{av} = P_a P_v + P_a (1 - P_v) + P_v (1 - P_a), \quad (1)$$

which reduces to

$$P_{av} = P_a + P_v - P_a P_v, \quad (2)$$

where P_{av} is the probability of a detection using both senses while P_a and P_v are the probabilities of detection using the single senses of auditon and vision, respectively. The major drawback of this model is that it consistently overpredicts the bimodal detection performance (Craig, Colquhoun, and Corcoran, 1976).

In a model proposed by Corcoran and Weening (1969), the eye and ear are also assumed to be independent channels; however, Corcoran and Weening propose a proportionality model. According to this model, a signal will always be reported as being present when both channels agree. When there is a conflict, detection by only one channel, a detection will be reported on a proportion of these cases. The proportional weighting is seen to be determined by the strength of the evidence upon which each modality makes its decision (Craig, Colquhoun, and Corcoran, 1976). The predictions of this model (excepting the refinement that the observer may elect to ignore information from a source whose reliability is not sufficiently high) can be estimated from the descriptive equation

$$P_{av} = P_a P_v + P_a (1-P_v) \times \frac{P_a}{P_a + (1-P_v)} + P_v(1-P_a) \times \frac{P_v}{P_v + (1-P_a)}$$

(Craig, et al., 1976). (3)

This theory has been shown to fit certain detection data (Corcoran and Weening, 1969) in addition to recognition data.

The output information from the auditory and visual channels are thought to be qualitatively the same at an internal analog level and as such may be readily combined. It is therefore postulated that seeing and hearing are equivalent to a *double look* (Green and Swets, 1966) or a *double listen.* That is to say that a single decision is made on the basis of information integrated from the two systems rather than two decisions, one made by each system, which are statistically combined later. While this model appears to be a very efficient description of the operation, in reality, it proves to be too efficient and therefore overestimates the observer's performance in the combined sensory condition. The equation of this model is

$$(d'_{av})^2 = (d'_a)^2 + (d'_v)^2 , \tag{4}$$

where d'_{av} is the index of detectability (Green and Swets, 1966) of the dual mode condition and d'_a and d'_v correspond to the indices of the respective single modes. Green and Swet's theory of signal detectability and the detectability indices for the signal pairs used in this thesis have received rigorous treatment by Janota (1977) and Martin (1978) and will not be reiterated here.

In an attempt to gain more insight into the dual mode processes, Bernstein, Rose, and Ashe (1970), Kohfeld (1969), and Nickerson (1973) investigated dual mode processing through the study of reaction time. The earlier studies of this model predict a facilatory effect due to the presence of the redundant stimuli. Nickerson argues that the presence of redundant stimuli serves to increase the subject's preparedness to respond as seen in reduced latencies.

Nickerson claims that this model operates in essence as a cueing theory system, in which one modality serves to cue or alert the other to the presence of a signal. The detection decision is then based on the output of the cued modality. The sensitivity of the cued modality is unaffected by this alerting, which is assumed to affect only the response criterion. In this model, it is assumed that one modality consistently cues the other. The alerting modality *decides* that a signal is present and cues the other modality. The alerting of the second modality increases the likelihood of a detection by the cued modality thereby causing a reduction in the latter's criterion and in the latency to *detect*. However, when no signal is present, the reverse is true;

that is, the criterion of the cued modality is raised and the detection is less likely to occur.

From this model, it is implied then that the dual mode detection criterion (β) is shifted in the direction of that of the criterion of the cueing modality, while the efficiency or d' remains unchanged. Facilitation will occur if the cueing modality has a lower criterion than that of the cued modality. The model states that dual mode efficiency is determined by the efficiency of the alerted modality. This appears to be true since from other studies observer's dual mode performance has been shown to be at least as good as the better of the individual modes and invariably superior to the poorer mode (Colquhoun, 1975; Loveless et al.,1970). The model described above yields the following equations, where β is the criterion to respond. Assuming the auditory mode to be more efficient than the visual, the equations read

$$|\beta_{av}-\beta_v| \leqslant |\beta_{av}-\beta_v|\text{, when } \beta_a \neq \beta_v \tag{5}$$

$$d'_{av} = d'_a \text{ and } d'_{av} \neq d'_v\text{, when } d'_a \neq d'_v \tag{6}$$

If the visual mode is more efficient, then the a and v subscripts will be interchanged.

In another model, the dual mode condition is assumed to be an input-output function (Craig, Colquhoun, and Corcoran, 1976). In this input-output model, the inputs and outputs from the two channels are seen as correlated variables, whose values are determined in the unimodal conditions. The justification for this model comes from Colquhoun (1975) who found a noticeable tendency for good visual detectors to also be good auditory detectors. Colquhoun found a high positive correlation for grouped data of observers in their performance on visual and auditory detection tasks. From this evidence, Colquhoun rejects the idea of the auditory and visual channels operating as independent information channels. The failure to find statistical independence between the two channels leads to the conclusion that, if there is a central decision-maker in the dual mode condition, then its inherent bias could account for the association between the two systems. That is, the overall performance of the dual mode system will be shifted toward that of the better of the two single systems. Craig et al. (1976) argue that lack of statistical independence could account for the overestimation in the predictions of the statistical summation and integration models. They propose the following equation as predictive of this model;

$$P_{av} = P_a + P_v - P_a P_v - \phi\sqrt{P_a(1-P_a)\,P_v(1-P_v)}\ , \qquad (7)$$

where (ϕ) is the correlation coefficient indexing the association between the auditory and visual inputs. It is assumed that ϕ ranges between 0 and +1 with a mean of 0.5. This equation resembles that of the statistical summation model with the only difference being the extreme right-hand term. Setting $\phi = 0$ yields Equation (1). Both of these models agree that a signal will be reported when either or both modalities indicate its presence. They differ in the extent to which they predict agreement rather than conflict between the sense modalities.

Clearly, the area of dual mode presentation has received a great deal of attention. Many of these studies have approached the problem from different directions, but all draw the same conclusions; dual mode presentation is superior to single mode presentation, but not as the sum of the two systems.

3.3 A Model of the Discrimination Process

It has been hypothesized that the way in which observers detect, recognize, and discriminate signals is through a method of feature extraction. That is, an observer has the ability to take in component stimuli, filter out extraneous information, and isolate the information which is relevant to the task at hand. For the discrimination task, the observer can compare the features of the stimuli most recently presented with a representation of the component features from another stimulus which has been stored in memory. The most recently occurring stimuli, after the irrelevant information has been stripped from it, may be thought of as a perceptual trace and the stored representation against which it is matched for discrimination may be thought of as a memory trace. The terms memory trace and perceptual trace have been used in a model of motor learning by Adams (1971), but the basic definition may also be applied here. Adams defines a memory trace as the choosing and initiation of a previously learned movement pattern which is stored in memory. However, this memory trace as defined by Adams is not equivalent to an engram but merely an active short duration neural unit. Adams defines a perceptual trace as the reference mechanism which uses proprioceptive and kinesthetic feedback to correct the movement process. In the case presented here, the memory trace like Adams' is stored in memory, but as the features of a signal pattern rather than a movement. The perceptual trace is contrasted to it and, based on same/different feedback the observer discriminates the two.

The amount of information an observer can extract from a broadband signal or feature is proportional to the amount of energy and the signal-to-noise ratio of the signal (Martin, 1978). Eriksen and Hake (1955) asked subjects to discriminate between visual stimuli differing in the dimensions of size, hue, and brightness. The stimuli were paired and differed in one, two, or all three dimensions. It was found that when the stimuli varied along two dimensions, discrimination was more accurate than when it varied along one dimension. When the stimuli differed along all three dimensions, discrimination performance was almost perfect. In general then, it appears that an ideal observer can make a discrimination based solely upon the most detectable feature characterizing the difference between the stimuli. For the auditory stimuli, the subject would detect the presence of the dichotomous feature by the amount of energy present in the feature. While for the visual stimuli, the subject would detect the dichotomous feature by detecting the presence of the peak of the octave band in question. That is, the subject would visually detect the peak of the octave band, for this would be the dimension along which the stimuli would differ.

For the experiment reported in this thesis, the signals used differ by only one feature; however, the model described could easily be extended by the addition of additional feature extractors to cover the multiple feature case.

The steps of stimulus perception in a detection, discrimination paradigm are:

1) Signal reception and initial encoding into neural pulses.

2) Feature extraction--extraction of relevant information from the input signal. These features are not necessarily identical to the acoustic and visual features characterizing the signal, but are correlated with these features (Reed, 1973).

3) Comparator stage--memory and perceptual traces are contrasted.

4) Alternative stage--respond or reprocess decision.

It is the feature extraction stage that we are interested in here. In the process of eliminating extraneous variables, the stimuli are reduced to a set of psychological dimensions or a perceptual trace which is compared to the memory trace. For the memory trace to be adequate for comparison, it must contain a feature list which is unique from all other patterns so that no confusion among patterns arises.

In the case where a signal pattern is detected, the pattern is then analyzed by the feature extractor. The feature extractor then reduces the signal to its component features. These features are symbolized as omega (ω). A proposed model of the discrimination process is shown in Figure 1 for the case where the discrimination between the perceptual trace and the memory trace is characterized as follows:

Signal 1--memory trace, contains features ω_1 and ω_2. Where ω_1 is the dichotomous feature and ω_2 is the fixed feature.

Signal 2--perceptual trace, contains feature ω_2 only. Therefore, the discrimination task involves discriminating between one signal pattern containing two features and a signal pattern that contains only the fixed feature. To an ideal observer, the fixed feature information is irrelevant when signal 1 is presented, although for real observers, the type of fixed feature has been shown to affect the hypothesis tests for the dichotomus features (Martin, 1978).

The model depicted in Figure 1 consists of the four previously mentioned steps of stimulus perception. The procedures which take place will be discussed in the following paragraphs. It must be borne in mind, however, that this model is not an attempt to accurately follow the steps of the human information processing system. It rather is a framework for predicting performance on the types of discrimination tasks illustrated. The model is seen to apply in both the auditory and visual discrimination tasks as well as in combined sensory tasks.

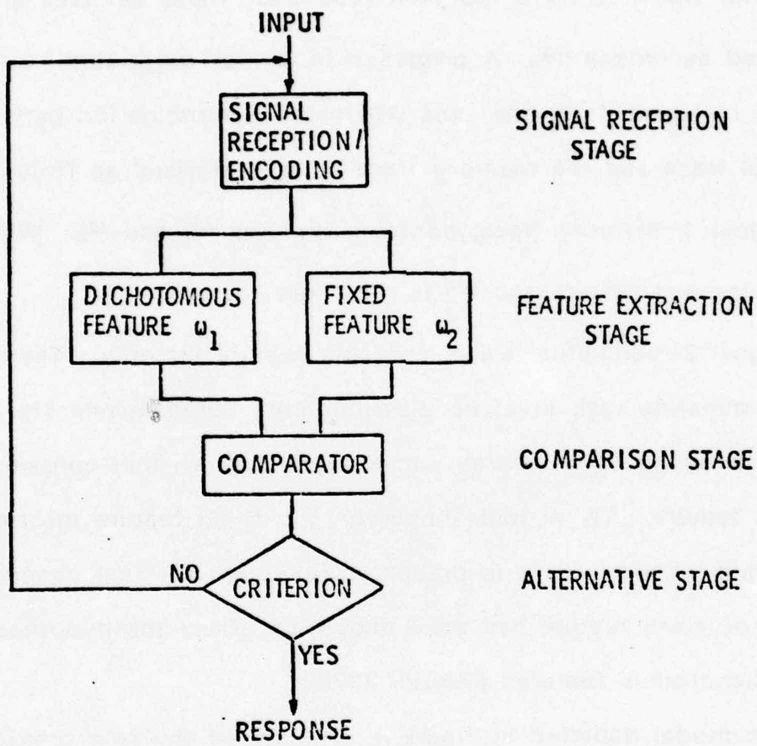

Figure 1. Information-processing model of the Feature Extraction Process.

The signal reception stage includes all of the physiological processes which occur in the human to convert both light energy and acoustic energy to neural energy which can be handled by the nervous system and the brain. These would include the bleaching of the photo receptors, stimulation of the cochlea, etc. These actions, however, are not of relevance to this thesis.

The output from the receptor stage serves as input to the feature extraction stage. In this stage, the observer is seen to break down the information into its component parts and to strip away unneeded or irrelevant information. In the auditory mode this is seen to be accomplished by a set of filters for detecting noise and envelope detectors to detect modulation (Martin, 1978). For vision, the same processes are also seen. That is, there are filters for detecting the spacial modulation of the light source and envelope detectors for detecting temporal modulation.

Within a specific modality, the system is seen as being able to extract more than one feature at a time. That is, the feature extraction process is seen as taking place in parallel following the pattern of Neisser's model (Reed, 1973). In the dual mode case, it is also possible that the feature extraction system is able to handle both auditory and visual information together in a multiplexing type network, or possibly there are separate channels for each. If each channel performs the process separately, then the outputs must be joined at a later point.

The next stage of the model is the comparator stage. Here, the perceptual and memory traces are compared. The terms memory and perceptual trace are arbitrary and only indicate which signal has been presented more recently. The most recent signal becomes the perceptual trace which is compared against a previously learned signal pattern stored as a neural trace in memory. The comparator contrasts the features of the two traces and draws a binary conclusion, same/different. The conclusion drawn is then passed on to the final stage, the Alternative stage.

In the Alternative stage, the observer takes the binary output from the comparator and based upon the same/different decision, either elects to respond or return to the input stage and repeat the process, comparing the perceptual trace to a different memory trace. If the comparator's criterion for same are satisfied, the observer responds. It must be stressed that the decision of the comparator is a binary. If there is any interaction of features in the extraction stage or if features are missed or irrelevant information included, the perceptual and memory traces will not match, the comparison criterion for same will be rejected, and either the process repeated or an erroneous response decision will be made.

The model discussed above describes the use of two signal patterns for each discrimination. The model assumes one dichotomous feature and one fixed feature. However, the model's application could easily be expanded to handle a greater number of both fixed and dichotomus features. However, the comparator stage would still only handle two signal patterns at a time, although other signals may be buffered, awaiting entry into the system.

3.4 The Hypothesis of Dual Sensory Presentation

Since the eye and ear are both capable of receiving energy independently, it would seem logical that higher order processing would also be independent. However, based on the findings of Colquhoun (1975), this may not be the case. As discussed earlier, if information from the two senses is input to the feature extraction model simultaneously, the model may not be able to process both information streams concurrently. The possibility of dual systems, one for each modality, exists. However, as discussed earlier, the outputs of these systems would somehow have to be joined statistically to allow for a response decision to be made.

Also discussed previously was the possiblity of a multiplexing system, whereby inputs from both systems could be processed in a time-sharing type system. This time-sharing system seems a viable idea based on the studies of reaction time using auditory and visual stimuli. It has long been known that reaction time to auditory stimuli is faster than reaction time to visual stimuli by about 50 msec (Sage, 1977). Sage goes on further to say that it has been shown that auditory stimuli reaches the cerebral cortex 8-9 msec after stimulation, while visual stimulation takes 20-40 msec before reaching the cortex. This difference in arrival time of stimulation in the dual presentation case may aid the processing in allowing for the processing of the auditory information before the arrival of the redundant visual information.

These differences in neural conduction times fit more than one of the previously discussed models. The cueing theory proposed by Nickerson (1973) states that one mode tends to cue or alert the second mode to the presence of a signal. It is quite possible that the more rapidly arriving auditory information serves to cue the visual sense in the redundant presentation case. The cueing theory implies that there are two channels, one for each sense, and that these channels work in parallel. As discussed earlier, if there are two channels, there is the need for a summing point where the output of each of the processors is combined to yield the respond decision. This combination of information may indeed be a statistical summation as proposed by Loveless, Brebner, and Hamilton (1970). That is, that following the comparator stage of the model there would be a combining of the two outputs of these separate channels before the alternative stage. In the alternative selection stage, the outputs would be contrasted and based on their degree of agreement, a response decision would be made.

A schematic diagram of the dual modality case is shown in Figure 2. This diagram depicts the dual mode processing, the inherent delay in the visual system, and the common summation point.

The hypothesis of this thesis states that, for certain types of signals, the combined use of both the auditory and visual systems will yield a higher detection rate and a better classification performance than the single mode case. The parameters used to assess these

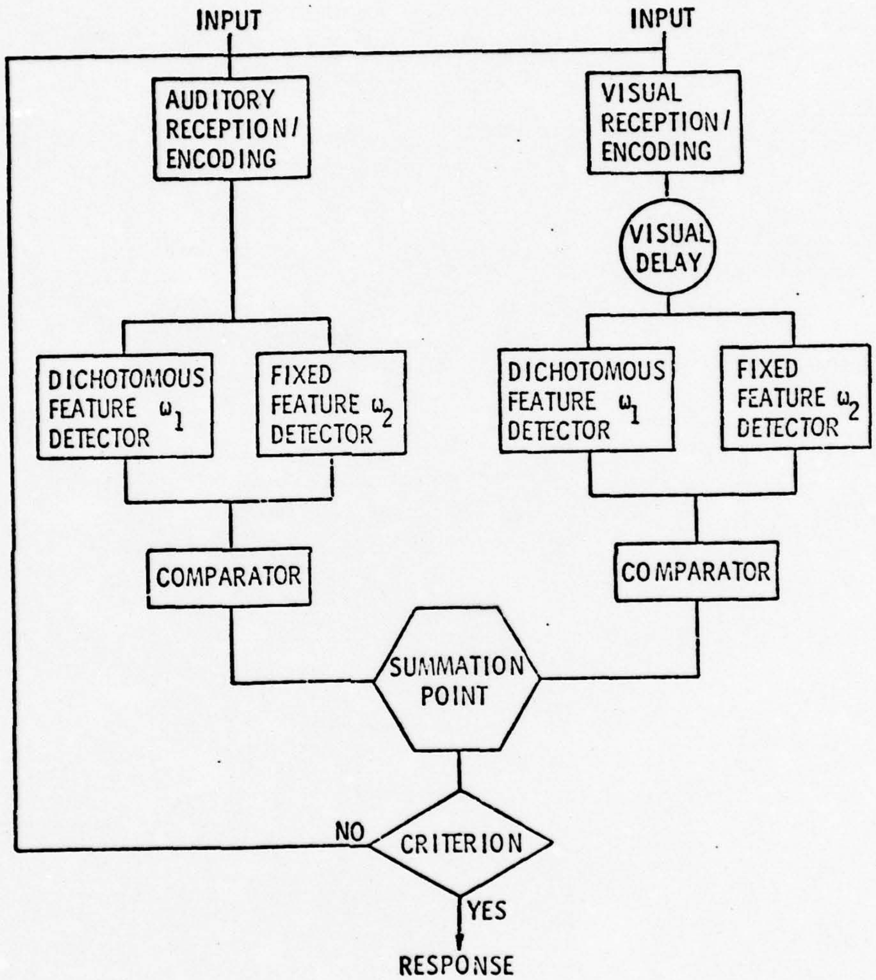

Figure 2. Dual modality information-processing model.

performances are the mean SNR at response and the probability of a correct response.

CHAPTER IV

METHOD

4.1 General

In this chapter, the experiment designed to assess the utility of using more than one sensory modality for presenting information will be discussed. The experiment was conducted at the Applied Research Laboratory of The Pennsylvania State University. The experiment was conducted over a four-month period using six pairs of laboratory-generated sounds as stimuli. The subjects were five university students, each of whom served for the duration of the experiment. The methods of data collection and reduction were automated as much as possible to ease the task of handling the large volumes of data that were generated.

Section 4.2 details the choice and construction of the stimuli used in this experiment. The signals used are composed of various fixed and dichotomous features.

The experimental procedures used to investigate the discrimination task is the modified threshold technique developed by Janota (1977). This procedure is discussed in Section 4.3. In the experiment, subjects were presented with two signals which differed by the presence or absence of the dichotomous feature. One of the signals was then presented against a white noise background. The SNR of the signal was initially very low and increased to allow the signal to emerge from the noise.

When the SNR had increased sufficiently, subjects were able to make a response under the criterion that they were *reasonably certain* of their choice as to which signal was presented. The same signal pairs were used for a group of six events with three such groups for each test session. Depending on the mode of the experiment, subjects were presented the stimuli either aurally via headphones, visually via an oscilloscope, or via both pieces of apparatus simultaneously.

Section 4.4 presents a description of the equipment used to generate and record the stimuli, construct the trials, and record the data. Section 4.5 discusses the procedures used for screening, selection, and training of the subjects used in the experiment. Section 4.6 details the methods of data analysis for the experiment. The measures of interest were the SNR at which the subject was willing to make a terminal decision as to which signal was being presented and the probability of a correct response *P(C)*. These variables are functions of stimulus complexity, and since the construction of the experiment allowed for different features of the signals to become detectable at different levels, the data should yield a trend of information content which will lend credence to the information processing aspects of the experiment.

The signal which is presented in the noise is called the probe stimulus. Data are only presented and analyzed for the cases in which the dichotomous feature is present in the probe stimulus since discrimination of signals in the feature absent case appears to use different information-processing techniques which are beyond the scope of this thesis.

4.2 Choice and Construction of Stimuli

In order to test the hypothesis detailed in the previous section and examine the feature analysis processes which are used in discrimination tasks, tests were conducted using three pair of signals. These signals were laboratory-generated sounds made up of octave bands of stationary noise at various frequencies and, in one case, amplitude modulation of a noise band by a 10 Hz square wave signal with a 50% duty cycle. The signals were generated using a General Radio GR-1390 Random Noise Generator, a Hewlett-Packard HP-3722 Noise Generator, Spectrum LH-42D and SKL band-pass filters, and several custom built components at the Applied Research Laboratory including a two-input mixer, a summing amplifier, and a square wave generator and modulator.

The background noise used was basically a white noise with frequencies below 70 Hz filtered out to avoid audio tape saturation (Janota, 1977). The 1/3-octave spectrum of this noise is shown in Figure 3. The background noise was produced by using a General Radio GR-1390 Random Noise Generator, whose level was adjusted by means of a General Radio stepped attenuator. The loudness levels of the test stimuli were controlled during the tests to be 65 phons (GD) (ISO R532). This loudness level was verified from the 1/3-octave band measurements of the voltage function to the headphones and taking into account the factory earphone calibration with the MX-41/AR cushions (Janota, 1977).

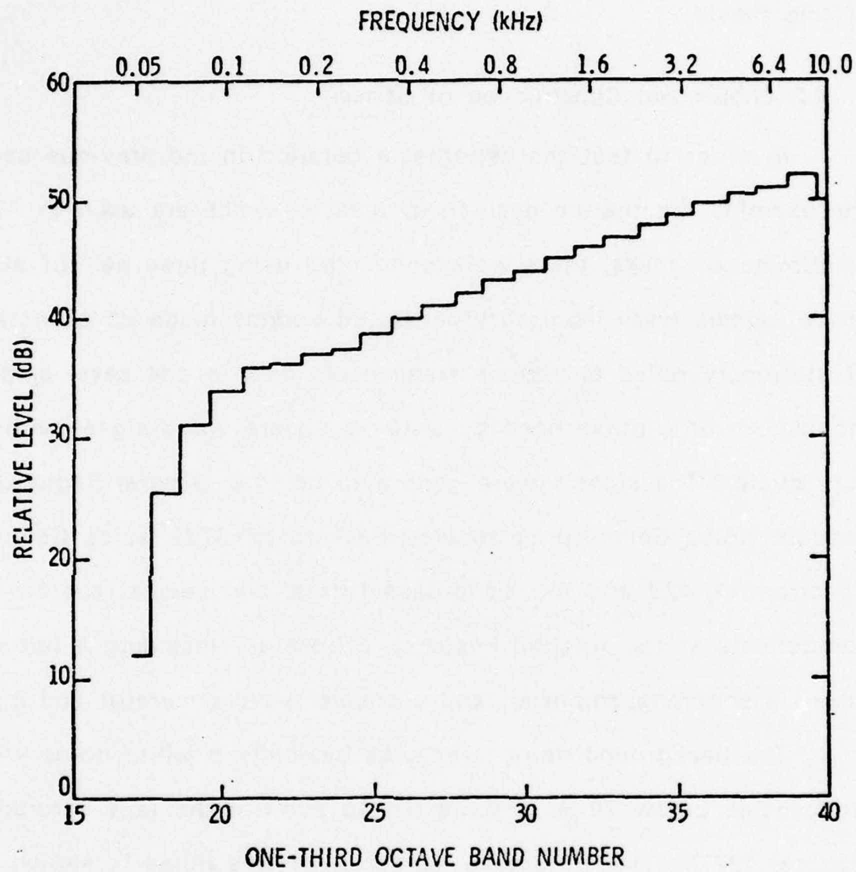

Figure 3. One-Third octave spectrum for background noise against which the signals were presented.

This was accomplished with a balanced mixer and a simple automatic loudness control built at the Applied Research Laboratory (Janota, 1977). The balanced mixer was used to change the SNR stepwise with time in 1/2-dB steps every 2 seconds, from an initial very low level where discrimination was not possible, to a much higher value where the tasks were considerably easier.

The signals were recorded and played back on a Crown 700 1/4-inch stereo tape recorder. The test stimuli were recorded on one channel, while the other channel contained electronic control signals for the apparatus. All of the test signal pairs had been used previously in auditory only discrimination studies by Janota (1977) and Martin (1978). These previous studies used the auditory mode only to assess the detection and discrimination abilities of human listeners.

The acoustic features composing these signals are listed in Table 1. The choice of signals used in this experiment was such that each signal pair contained one dichotomous feature and one fixed feature. The signals listed in Table 1 are paired, that is, signals 1 and 2 make up a signal pair, signals 3 and 4 make up a pair, etc. The signals were created so that features had equal energy in the bands; that is, the narrower bandwidth features had higher spectral levels. For the signal in which amplitude modulation was the feature, construction was such that for the pure signal without background noise, the ratio of $\Delta I/I$ in dB, was approximatly 0.6 measured in the modulated band. This corresponds to a Weber fraction of approximately -2 dB. The intensity

TABLE 1

DESCRIPTION OF FEATURES COMPRISING THE SIGNALS

SIGNAL	PATTERN	DESCRIPTION
1	.1	Octave band of stationary noise centered at 500 Hz. (band 27).
2	.1	Pattern .1 amplitude modulated by a 10 Hz. square wave.
3	.2	Octave band of stationary noise centered at 500 Hz. (band 27).
4	.2	Pattern .2 plus a band of stationary noise centered at 4000 Hz.
5	.3	Octave band of stationary noise centered at 250 Hz. (band 24).
6	.3	Pattern .3 plus a band of stationary noise centered at 1000 Hz.

increments were characterized by effective durations of 50 msec and bandwidths corresponding to bands centered at 500 and 4000 Hz octave bands. With these bandwidths, durations, and intensity ratios, the modulation was quite pronounced. The addition of the background noise to this signal effectively reduced the ratio of $\Delta I/I$ so that, at the initial starting point of the trial when the SNR was low, the modulation could not be perceived (Martin, 1978).

From Table 1, it can be seen that the signal pattern denoted .1 consisted of two signals both centered at 500 Hz, one using amplitude modulation as the dichotomous feature. The spectral plot of the signal used in this treatment can be seen in Figure 4. Since amplitude modulation cannot be seen in the visual display, only one member of the signal pair is shown. Spectral plots for signal patterns .2 and .3 can be seen in Figures 5 and 6, respectively. The designations $SH1$ and $SH0$ relate to the feature present and feature absent cases, respectively.

4.3 Experimental Design

The modified threshold technique, previously used by Janota (1977) and Martin (1978), was used here to obtain information about the performance on the three discrimination tasks discussed in the previous section. Each experimental trial consisted of an exposure set and a response period. During the exposure set, the signals were presented without interfering noise, the first signal in this random ordering was denoted *signal A* and the second signal was denoted *signal B*. These designations were purely arbitrary and were not indicative of signal

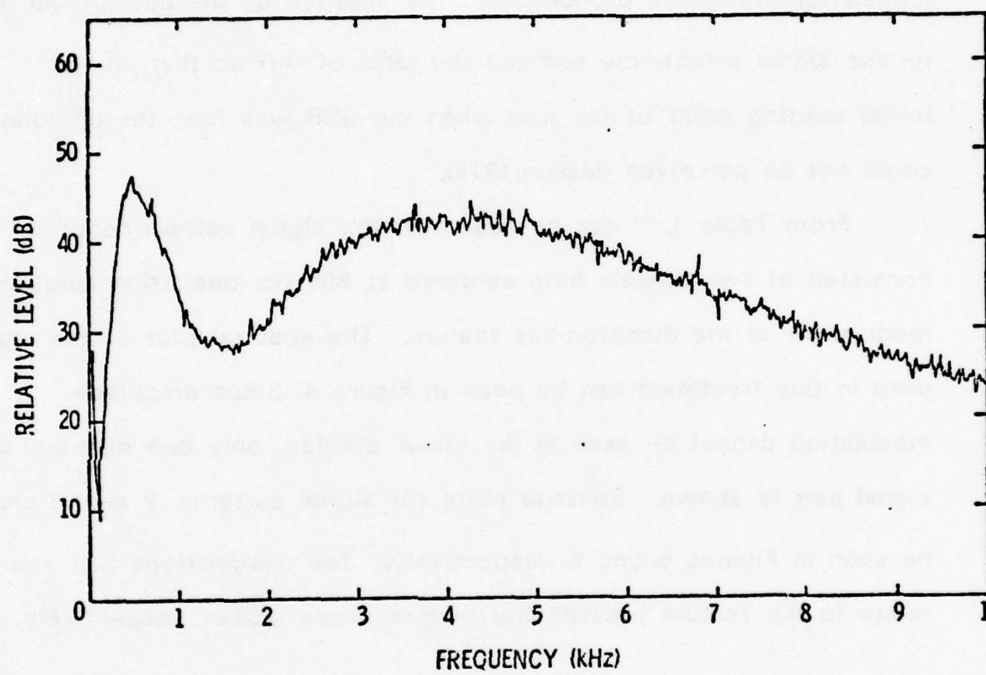

Figure 4. Spectral plot of signal pattern .1.

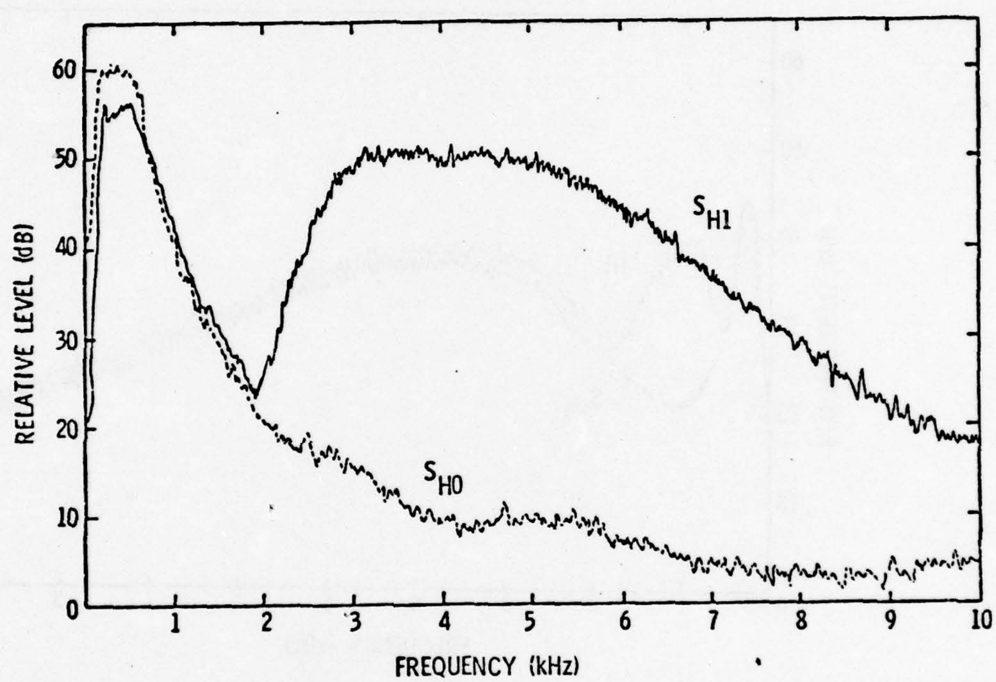

Figure 5. Spectral plot of signal pattern .2.

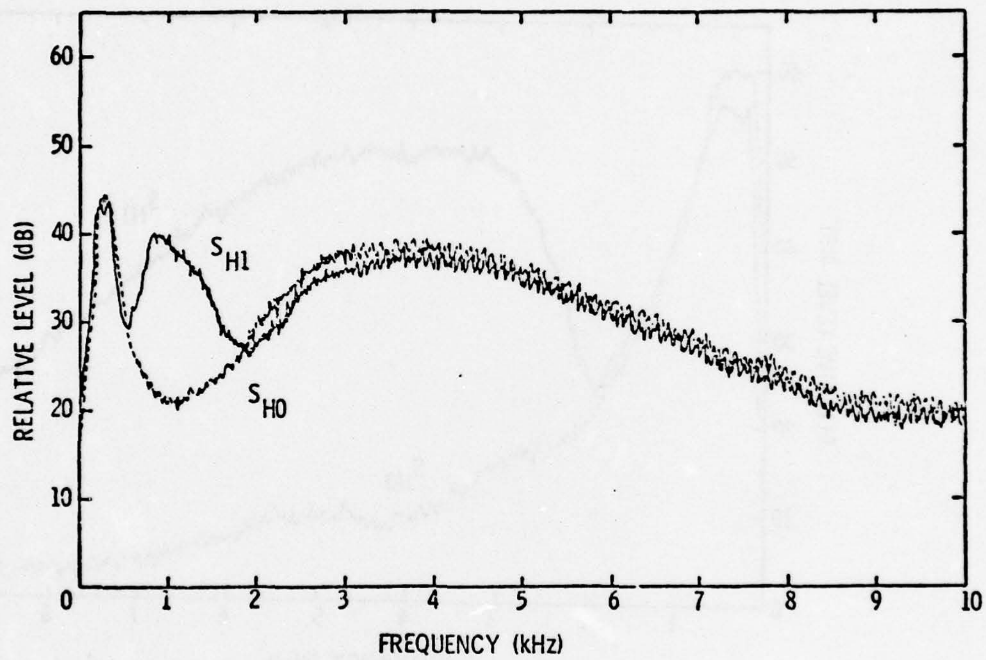

Figure 6. Spectral plot of signal pattern .3.

characteristics. During the response period, either *signal A* or *signal B* was presented against the background noise, the signals each had an equal probability of occurrance.

According to the procedures of the modified threshold technique, the probe signal, signal to be detected, was initially presented at a very low SNR, and then the SNR was increased stepwise by 1/2 dB every 2 seconds. The increases in signal occurred gradually and there were no transients to indicate the step change. Subjects were therefore unable to report when the steps occurred (Janota, 1977).

The starting SNR's of the signals were randomized, being chosen from a set of values ranging over 4 dB. This randomization was done in an effort to avoid the possibility of subjects responding due to elapsed time rather than at a specific SNR. The total time of each event was also ramdomized over a period ranging from 56 seconds to 1 minute 15 seconds. However, as will be discussed later, these efforts appear not to have been completely successful.

At the beginning of the response period, the probe signal was completely masked by the background noise. As the trial progressed, the SNR would increase until either the subject was willing to respond or until the time for the event was exhausted. The subjects responded by pressing one of two designated response buttons. The buttons were denoted A and B, respectively, and these designations corresponded to the order in which the signals were presented in the exposure set. The subject's response was electronically recorded on a cassette tape; this recorded response contained information as to the SNR at which the response was made and also the elapsed time of the event from the beginning of the event to the point of the response. At the response, the signal was blanked so that the subject received no feedback as to the correctness of the choice. The primary reason for not supplying knowledge of results to the subject was due to the design of the hardware used.

The signals used as test stimuli were recorded on a Crown 700 tape recorder with 18 trials per 45-minute session. The degree of automation of the test apparatus allowed the subjects to run test sessions at their convenience with the only constraint being that no two sessions be run sequentially and that not more than two sessions be run in any 24-hour period. The test apparatus was such that conversion from one sense modality test to another could be accomplished with a minimum of modification of the apparatus. All of the test sessions were conducted in an audiometric booth.

During mode 1, the auditory portion of the experiment, subjects were presented the signals via calibrated Telephonics TDH-39 headphones. The test booth was small but comfortable and contained a chair, a shelf which supported the response recorder, a wall-mounted lamp, and a window. The Crown recorder and cassette recorder were located outside the booth and signals were fed through a patch panel to the subject. To conduct a test session, the subject would mount a designated tape on the Crown, load a response cassette, and enter the test booth; the session would last approximately 45 minutes. At the close of the test session, the subject would rewind the test tape, dismount it, and replace the tape in an assigned rack. The subject would then fast forward the cassette and place it in an assigned folder.

The experimental trials were recorded in blocks of six events with the three different signal pairs comprising an 18 event test session. The three signal pair treatments were presented in each session and each experimental task could be termed as *easy, medium,* and *hard* tasks.

The modified threshold technique differs from classical signal detection theory techniques in that the signals used are bands of noise as opposed to a single frequency tone, signals are presented on each and every trial as opposed to the use of signal trials and noise trials, and the subjects respond during the trial as opposed to after the trial concludes. The signals used in the trials increase in SNR with time until the signal can be detected against the noise and the subject is willing to commit himself to a terminal decision as to which of a previously learned pair of signals is being presented. The subject responds under the criterion of *reasonably certain* and, upon response, the signal is blanked to eliminate the possibility of knowledge of results.

4.4 The Test Apparatus

The test apparatus consisted of a Crown Model 700 1/4-inch stereo tape recorder. This unit was used to play the test tapes which contained the recorded test signals on one channel and electronic control signals for the apparatus on the other channel. The audio channel which contained the stimuli also contained verbal instructions to the subjects. At the beginning of each test tape, there were instructions to the subjects explaining the test procedures and the methods of responding. In addition to these instructions, there were cut designations between events on the tape which would inform the subject which event just concluded and which event was about to begin. The test treatments were arranged in groups of six and the subjects were alerted as to when the treatment was about to change. A full description of the number of test tapes and their organization will be given in later paragraphs.

Integral with the Crown recorder was also a Sony Model TC-95L cassette recorder. This recorder was used to record the subject's responses. The responses were recorded on 90-minute cassette tapes with each side of the cassette being used for one test session. As mentioned previously, the test signals were patched through a patch panel to the inside of the audiometric booth. Inside the test booth, the signals were presented to the subjects via a pair of headphones which were calibrated to manufacturers specifications. Also inside the test enclosure was a response recorder with which the subjects made their responses. The response recorder consisted of two push buttons designated A and B to correspond with the two signals presented in each exposure set. Also on the response recorder was a green light to indicate the response period and two red lights denoted A and B to correspond to the response keys. During the response period, the green light was illuminated until the subject made a response, at which time the green light was extinguished and one of the red lights corresponding to the button depressed became illuminated. The subject's responses were recorded as a tone on the cassette. There were two different frequency tones denoting either an *A* or a *B* response.

Also used in the test apparatus was a Federal Scientific UA-500 Ubiquitus Spectrum Analyzer and a Hewlett-Packard 5-inch Model 122AR oscilloscope with P1 short persistence phospher. The output of the Crown recorder was fed through the spectrum analyzer and displayed as the instantaneous specrta of both the signal and the noise on the oscilloscope. The spectrum analyzer and oscilloscope were only used by the subjects during the test sessions in modes 2 and 3, the visual and audio-visual portions of the experiment. During mode 1, the auditory portion of the experiment, the visual display was turned off. The visual display system, the oscilloscope and spectrum analyzer, were placed on a table outside the test booth and only the oscilloscope screen was visible by the subject through a window in the side of the booth. The visual display was positioned to be at approximately eye level for the seated subject and at a distance of less than 3 feet. The intensity of the trace on the oscilloscope was set for comfortable viewing, and the background illumination in the booth was 6.15 foot candles, as measured by a Spectra Brightness Spotmeter, Model 1415-UB. This background illumination was provided by the wall-mounted lamp, as the window was shrouded except for the visual display screen.

The test stimuli consisted of a set of six prerecorded quarter-inch audio tapes. Each of these tapes contained each of the treatment signal pairs clustered in groups of six events. Each of the experimental treatments were used in random order on each tape. The instructional set that was included at the beginning of each tape can be seen in Appendix A.

4.5 Subject Selection and Training

The subjects who took part in this experiment consisted of five university students; there were three males and two females, although sex was not a factor. The subjects were screened for normal hearing by recent audiogram measurements taken by The Pennsylvania State University Speech and Hearing Clinic. All subjects were also tested for nearfield visual acuity, either corrected or uncorrected, by the experimenter using a Titmus Vision Tester. Subjects were chosen who had no prior experience with psychoacoustic or signal detection experiments. Therefore, the signal patterns used were new and novel stimuli to all subjects.

Upon agreement to participate in the experiment, the procedures to be followed were fully explained to the subjects. The subjects were also asked to sign an informed consent form and told that they may cease their participation in the experiment at any time. Copies of the informed consent form and instructions to subjects are shown in Appendix B.

The degree of automation of the test apparatus allowed the subjects to conduct experimental sessions at their convenience, and the experimenter was only present at the first two session of mode 1 for each subject and also at the first session of modes 2 and 3, in order to instruct the subjects how to convert the test apparatus from one mode to another. Following these sessions, the subjects were quite able to handle all aspects of the test apparatus with no difficulty.

In previous studies, Janota (1977) and Martin (1978) using the same and similar test stimuli and the same procedures showed that, for naive subjects, performance changed dramatically over the first five sessions, but stabilized thereafter. Based on these findings the data for the first five sessions in each mode were not analyzed. Cornell (1978) has shown that subjects trained in the modified threshold technique show very high consistency when data collected for a given signal pair are compared between early and late sessions.

4.6 Methods of Data Analysis

The methods of data analysis for the experiment are dicussed in this section. The determination of which pieces of data may be pooled, which may be eliminated, and what statistical tests would be most meaningful and valid will be addressed. Martin (1978) states that subjects indicated that a sample of the signals in noise would have been helpful prior to the first event of a group and that they used the first event as just such a sample. Janota (1977) had also observed this fact. Therefore, the data for the first event in each group have been deleted from the data set.

The quantities of interest on each event are the classification decision, SH1 or SH0, the signal-to-noise ratio at response, and the elapsed time from the start of the event to the response. After the appropriate data are grouped, the items of interest are the probability of a correct classification and the mean and standard deviation of the SNR at response.

On each trial, the subject had to make a classification decision about the signal. These decisions taken over a large number of trials allow for the determination of the probability of a correct response, *P(C)*, for the particular signal group. Assuming that the events comprise Bernouli trials with equal probability of occurrence, an approximately Gaussian distribution can be obtained for X observed correct classifications in N trials (Janota, 1977). Using the appropriate transform, the 90% confidence limits on *P(C)* are given by:

$$\left\{ \mathrm{Sin}\,\left[(2 \text{ arc Sin} \sqrt{X/N}) - 1.69/\sqrt{2N}\,\right] \right\}^2 \leqslant P(C) \leqslant \left\{ \mathrm{Sin}\,\left[(2 \text{ arc Sin} \sqrt{X/N}) + 1.69/\sqrt{2N}\,\right] \right\}^2 \quad \text{(Janota, 1977)} \quad (8)$$

To allow the confidence interval to be sufficiently narrow for statistical significance, a large number of events are required. To express *P(C)* to within ± 10% of its actual value, fifty to seventy events are required (Janota, 1977).

The major point of interest in this experiment is the SNR where the subject is able to make the classification decision. This SNR is the level of the signal above the noise at the point where the response is made. That is, the SNR is the average level in dB of the signal relative to the noise in the dichotomous band. The SNR is calculated as the sum of the 1/3-octave-band levels in the dichotomous portion of the signal minus the sum of the 1/3-octave-band noise levels in the same bands (Martin, 1978).

The SNR for each dichotomous feature is given by the equation

$$SNR = l_o + l_b \qquad \text{(Martin, 1978)} \qquad (9)$$

Where l_o is the average level of the feature above the noise at 0 dB balanced mixer setting, and l_b is the balanced mixer setting corresponding to the point of response and corrected for mixer nonlinearity (Martin, 1978). After determining the SNR for each dichotomous feature on each trial, the data from similar populations may be pooled. From these data, the associated first order statistics may be obtained, that is, the mean SNR and standard deviation. Janota (1977) has shown that the distribution of response SNR'S may be regarded as Gaussian, given a small number of no-response events. Janota goes on further to state that from this relationship, the 90% confidence interval for the means is given by Equation 10.

$$[\bar{X} - t_{\alpha/2}(N-1)\frac{\sigma}{\sqrt{N}}] \leq \mu \leq [\bar{X} + t_{\alpha/2}(N-1)\frac{\sigma}{\sqrt{N}}]$$

(Freund, 1971) (10)

where $\alpha/2$ is the confidence interval for the t-test.

The third measure, response time, is unfortunately highly correlated with the response SNR. Although steps were taken to try to separate these two parameters, the attempts did not prove to be successful as will be discussed in Chapter V.

To obtain the data required for calculation of the aforementioned parameters, decisions must be made concerning the data. That is, it must be decided which data to include, which data to omit, and which data may be pooled. These decisions for the data reported in this thesis are as follows:

1) All data in which the feature absent case was presented as the probe in the noise have been omitted. As stated earlier, discrimination of these types of signals involves processes which are beyond the scope of this thesis.

2) The data for each subject from the first five sessions of each mode have been omitted. These sessions were considered as training sessions; therefore, their data have been omitted.

3) Data from the first event of each group of six were omitted. Subjects tended to use this event as a sample of the signal in the noise. Janota (1977) and Martin (1978) have shown significant differences in performance when this event is compared to the remaining five events in the group.

4) Data from all events where subjects failed to respond. Omission of these relatively few cases did not result in a significant shift in the distribution of the responses.

Following these omissions, the data were pooled for each treatment. The rationale for this is given by Cornell (1978) who found between-subject variablity was comparable to within-subject variability for subjects trained in the modified threshold technique. Therefore, data for all five subjects have been pooled for subsequent analysis.

CHAPTER V

RESULTS AND DISCUSSION

5.1 General

The results of the experiments with the three signal pairs will be presented in this chapter. The procedures to be followed in the analysis will follow those noted in Section 4.7. The results of the experiments will be discussed in terms of the feature extraction model outlined in Section 3.3. It will be shown that the data do not support the original hypothosis of this thesis; however, a second order effect was found to be supported.

Table 2 depicts a summary of the data of the measures of interest in the experiment. Column 3 yields the number of valid events upon which the statistical tests of each mode are based. The number of data points listed in column 3 are based on the procedures outlined in Section 4.7. In column 4 is shown the mean SNR, level of the feature above the noise, at which the subjects were willing to respond. Column 5 yields the sample standard deviation of the SNR, and column 6 lists the observed probability of correct responses. The remainder of this chapter will detail the analyses and interpret the data listed in Table 2. The signal pattern entries in Table 2, column 2 denote the mode of presentation and signal pair. The integer portion of the entry reflects the mode of presentation with 1. being the auditory only, 2. being the visual only,

TABLE 2

DESCRIPTIVE STATISTICS OF THE RESULTS OF THE SIGNAL PRESENTATION MODES

EXPERIMENTAL MODE	SIGNAL PATTERN	N	MEAN SNR	STANDARD DEVIATION OF SNR	P(C)
1	1.1	46	3.34	3.28	0.8214
1	1.2	33	7.88	5.27	0.7333
1	1.3	23	7.17	4.95	0.6571
2	2.1	22	4.22	5.94	0.5116
2	2.2	43	5.56	2.96	0.9772
2	2.3	32	6.59	3.40	1.000
3	3.1	53	4.68	2.98	0.9464
3	3.2	45	6.01	1.96	1.00
3	3.3	33	6.50	2.33	0.9428

and 3. being the combined audio-visual. The decimal portion of the entry lists the signal pair, with .1 being the amplitude modulation case, .2 being the band of noise centered at 4000 Hz as the dichotomous feature, and .3 denoting the signal pair where the band of noise centered at 1000 Hz is the dichotomous feature.

Section 5.2 details the analysis of the data for mode 1, the auditory-only portion on the experiment. The signal patterns will be contrasted as to the listener's performance in the detection and discrimination tasks. The signal patterns will be discussed in the order of presentation in Table 2. Section 5.3 will detail the data for mode 2, the visual-only portion of the experiment, and will compare the listener's performance in this mode to that in mode 1. Section 5.4 will contrast and compare the findings in mode 3 to both modes 1 and 2.

5.2 Results of the Auditory-only Case

The data presented in this section will follow the order of presentation listed in Table 3. Signal patterns 1.1, 1.2, and 1.3 will be discussed and contrasted in terms of their features and the subject's performance in detecting and discriminating these signals. Table 3 presents a summary table of this presentation mode and details the number of correct responses out of the number of possible responses, the calculated probability of a correct response, and the 90% confidence interval for the probability correct.

TABLE 3

SUMMARY TABLE OF SIGNAL PATTERN PERFORMANCE FOR AUDITORY PRESENTATION

SIGNAL PATTERN	CORRECT RESPONSES	P(C)	90% CONFIDENCE INTERVAL
1.1	46/56	0.8214	0.7170 - 0.9059
1.2	33/45	0.7333	0.5950 - 0.8515
1.3	23/35	0.6571	0.4836 - 0.8113

The mean SNR at response and the sample standard deviation are shown in Table 2. For signal pattern 1.1, the subjects displayed a slight degree of variability in response. Using amplitude modulation as the dichotomous feature, this signal was considered as easy to detect in the auditory mode, as illustrated by the *P(C)*.

For signal pattern 1.2, the band of noise centered at 4000 Hz, the subjects had difficulty in detecting and discriminating the signal. Table 3 also lists the pertinent infomation for this signal pattern. The mean SNR at response and sample standard deviation can be seen in Table 2. The subjects showed a great deal of variability in their responses and the signal pattern could be termed as difficult to detect.

Signal pattern 1.3, the band of noise centered at 1000 Hz, also presented difficulty to the subjects in detecting and discriminating the signals. Table 2 lists the SNR at response and the sample standard deviation for this signal treatment. Table 3 details the *P(C)* and its associated 90% confidence interval.

From examination of Table 2, it can be seen that signal pattern 1.1 yielded the highest detection and discrimination performance in the auditory-only presentation mode. This signal pattern will be discussed further in later sections of this chapter.

5.3 Results of the Visual-Only Case

The data presented in this section will detail the signal patterns denoted 2.1, 2.2, and 2.3. Table 4 presents a summary of this presentation mode and details the major points of interest.

For signal pattern 2.1, the amplitude modulation pattern with visual only presentation, the subjects encountered a great deal of difficulty in discriminating the signals. The SNR and sample standard deviation are shown in Table 2. Table 4 lists the *P(C)* and 90% confidence interval for this signal pattern.

For signal pattern 2.2, the 4000 Hz band in the visual-only presentation mode, the subjects had little difficulty in detecting and discriminating the signal pairs. The SNR at response and the sample standard deviation can be seen in Table 2. The *P(C)* and 90% confidence interval can be seen in Table 4.

The signal pattern 2.3, which used the 1000 Hz band as the dichotomous feature presented the subjects with a moderate amount of difficulty in detecting and discriminating the signals. Table 2 lists the SNR at response and the sample standard deviation for this case. Table 4 lists the *P(C)* and the 90% confidence interval.

Examination of Table 2 shows that signal pattern 2.3 yielded the best discrimination and detection performance for that signal pattern; this is shown by the *P(C)*. Possible reasons for this showing in performance will be discussed later in the chapter.

TABLE 4

SUMMARY TABLE OF SIGNAL PATTERN PERFORMANCE FOR VISUAL PRESENTATION

SIGNAL PATTERN	CORRECT RESPONSES	P(C)	90% CONFIDENCE INTERVAL
2.1	22/43	0.5116	0.3346 - 0.6871
2.2	43/44	0.9772	0.9235 - 0.9994
2.3	32/32	1.00	0.9778 - 1.00

5.4 Results of the Audio-Visual Case

This section will detail the combined, simultaneous use of both the auditory and visual presentation modes. The data presented in this section will detail signal patterns 3.1, 3.2, and 3.3. For signal pattern 3.1, the subjects had little difficulty in detecting and discriminating the signal patterns. The SNR at response and the sample standard deviation can be seen in Table 2 and the *P(C)* and 90% confidence interval are shown in Table 5. This signal pattern will be discussed in greater detail in Chapter VI.

Signal pattern 3.2, which uses the 4000 Hz band of noise as the dichotomous feature, provided little difficulty for the subjects in detecting and discriminating the signals. Table 2 shows the mean SNR and sample standard deviation for this signal pair, while Table 5 lists the *P(C)* and the 90% confidence interval.

Signal pattern 3.3, the 1000 Hz band, offered little difficulty to the subjects in detection and discrimination of the signal pattern. Table 2 displays the SNR at response and the sample standard deviation for this signal pattern, while Table 5 lists the *P(C)* and confidence interval.

5.5 Comparisons of Signal Patterns Between Modes

For signal patterns 1.1 and 2.1, the amplitude modulation case for the auditory and visual only cases, the t-test for difference of means showed no significant difference in the mean SNR to respond. However, Satterthwaite's F' shows the variances to be heterogeneous, $F'(42,55) = 3.28$, $p < .01$. In comparison, the subjects' responses showed greater variability in pattern 2.1 than in pattern 1.1.

TABLE 5

SUMMARY TABLE OF SIGNAL PATTERN PERFORMANCE FOR AUDIO-VISUAL PRESENTATION.

SIGNAL PATTERN	CORRECT RESPONSES	P(C)	90% CONFIDENCE INTERVAL
3.1	53/56	0.9464	0.8826 - 0.9862
3.2	45/45	1.00	0.9842 - 1.00
3.3	33/35	0.9428	0.8542 - 0.9911

The t-test for the difference of means for signal pattern 2.1 and 3.1 showed no significant difference in the mean SNR to respond; this finding was surprising and will be discussed more fully in Chapter VI. The sample variances yield a significant difference, $F'(42,55) = 3.98$, $p < .01$. The interpretation of this finding is that the subjects reduced their variability in going from mode 2 to mode 3. The subjects' responses in mode 3 were much less variable than in mode 2.

A t-test between signal 1.1 and 3.1 yielded a significant difference in means, $t(110) = -2.162$, $p < .05$. This would indicate the subjects' performance on signal pattern 1.1 was discernably better than their performance on pattern 3.1. The test of homogeneity of variance yielded no significant difference in variability between the two cases. Figure 7 presents a schematic representation of the subjects' response variability for signal pattern .1 across the three test conditions.

A t-test between signal patterns 1.2 and 2.2 yields a significant difference in means, $t(70) = 2.5673$ $p < .05$. The test for equality of variance showed the variances to be significantly different, $F'(44,43) = 3.15$, $p < .05$. Between modes 1 and 2 for this signal pattern, the subjects showed a marked reduction in variability in their responses.

For signal patterns 2.2 and 3.2, the t-test for difference of means showed no significant difference in the mean SNR's at response. However, the variances were found to be significantly different, $F'(44,43) = 2.29$, $p < .01$. This indicates that the subjects were less variable in their responses in mode 3 than in mode 2.

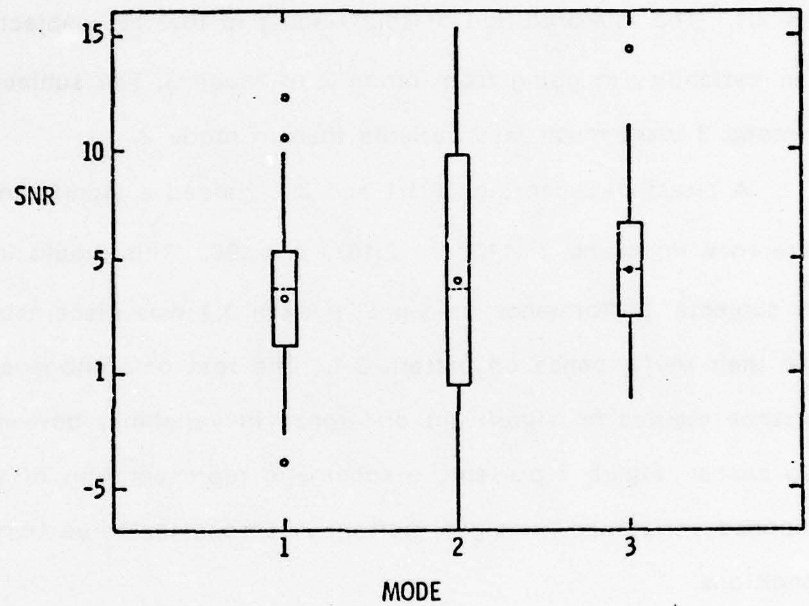

Figure 7. Variability in SNR for signal pattern .1 by presentation mode.

In comparing patterns 1.2 to patterns 3.2, the *t*-test for difference of means showed a significant difference in the mean SNR's to respond, t (56) = 2.2234, $p < .05$. This finding indicates that the subjects were able to respond at a significantly lower SNR in mode 3 than in mode 1. Also, for these signal patterns, the variances were shown to be significantly different, F' (44,44) = 7.23, $p < .01$. This indicates the subjects' responses were less variable in mode 3 than in mode 1. Figure 8 presents a schematic representaion of the subjects response variability for signal pattern .2 across the three test conditions.

For signal patterns 1.3 and 2.3, no differences in the mean SNR to respond were found. However, the variances did prove to be significantly different, F' (31,34) = 2.12, $p < .05$. The subjects' responses are less variable for signals 2.3 than for 1.3.

Comparison of signal pattern 2.3 to 3.3 show the means not to differ; however, the variances are significantly different, F' (34,31) = 2.13, $p < .05$. The subjects were able to decrease their response variability in mode 3 as opposed to mode 2.

In comparing signal pattern 1.3 to 3.3, again the means do not differ; however, the variances are significantly different. The comparison of variances yielded an F' (34,34) = 4.51, $p < .01$. This shows that subjects were more consistent in their responses in mode 3 than in mode 1. Figure 9 depicts a schematic representation of the subjects' response variability for signal pattern .3 across the three test conditions.

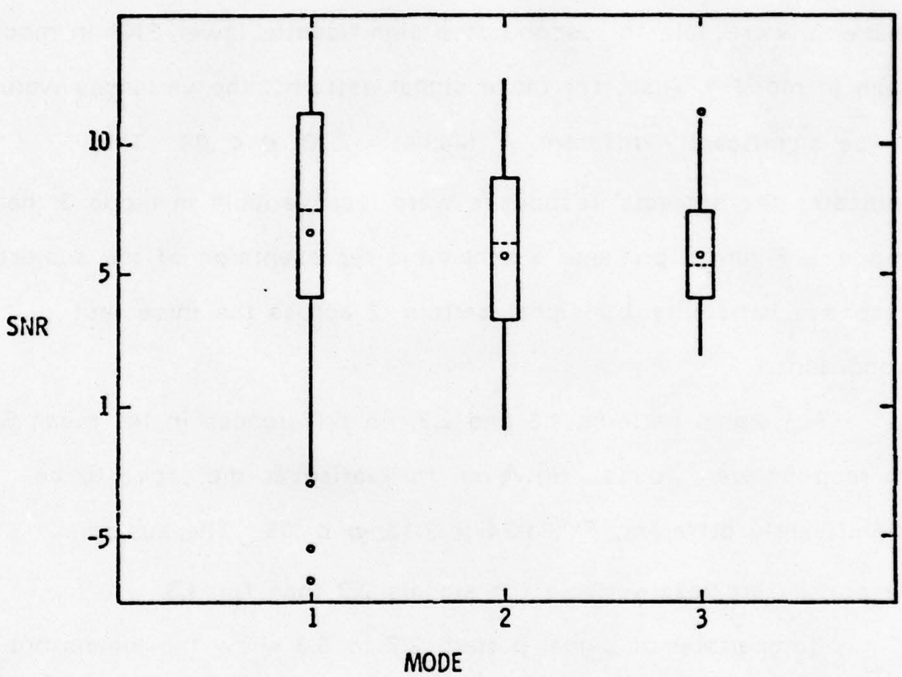

Figure 8. Variability in SNR for signal pattern .2 by presentation mode.

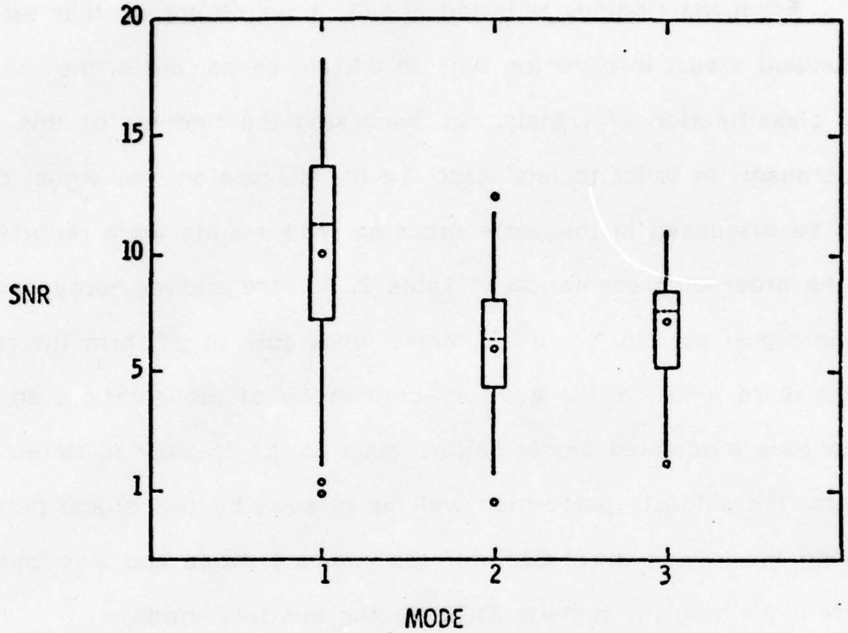

Figure 9. Variability in SNR for signal pattern .3 by presentation mode.

5.6 Discussion

From the findings reported above, it would appear that the use of redundant visual information will, in certain cases, aid in the detection and classification of signals. In discussing the findings of this experiment, in order to lend order to the discussion, the signal patterns will be discussed in the same order as their results were reported, which is the order of presentation in Table 2. For the auditory-only case, using signal pattern 1.1, the subjects were able to perform the required tasks quite well. In the auditory-only mode of presentation, an amplitude modulated signal pattern such as 1.1 is easy to detect and, hence, the subjects performed well as is seen by this signal pattern having the lowest mean SNR for the auditory mode and also by the signal pattern yielding the highest *P(C)* for the auditory mode.

In contrast to this, if we look at signal pattern 2.1 which is the amplitude-modulated signal presented visually, we find this combination to yield poor subject performance. The wide degree of variability in SNR indicates this fact since, in the visual presentation mode, amplitude modulation cannot be detected. Therefore, the signal pattern displayed on the visual display yielded no information to the subject as to the presence or absence of the dichotomous feature. The feature absent case appeared identical to the feature present case on the display. Since the two possible signals appeared identical, it was originally hypothesized that the subjects' performance in this mode would show a marked degradation in discrimination of the signals. This was not found to be the case. Although the subjects showed a significant increase in variability between modes 1 and 2, the mean SNR's did not differ. The data show the subjects to be responding at essentially the same SNR.

As mentioned previously, the starting SNR and the total elapsed time for each event was randomized in an effort to prevent the subjects from adopting a strategy of responding with time rather than SNR. These efforts appear not to have been successful since subjects do appear to be responding with time. It was hypothesized that the subjects' performance would degrade due to the lack of discriminability of the signal patterns; it was felt that this degradation of performance would manifest itself in a significantly higher SNR to respond. However, this was not the case since the t-test comparison with mode 1 showed no difference in the mean SNR at response. The breakdown of the subjects' ability to discriminate the signal is shown in the $P(C)$. It was hypothesized that the subjects would be operating at near chance level, which was borne out in the calculated $P(C)$.

Signal pattern 3.1, the amplitude modulation case with audio-visual presentation, will be discussed here. For this signal pattern, the amount of response variability was greatly reduced compared to the previous mode. However, the mean SNR to respond in this case was significantly higher than in mode 1. It was assumed here that since the auditory mode is yielding the greatest amount of information, then the subjects' detection and discrimination would be based totally on the auditory information. This appears to be the case; however, the presence of the visual display seems to act as a distractor to the subject as shown by the significantly higher response SNR over mode 1. That is, the subjects tend to detect the signals aurally, but monitor the visual display awaiting confirmation of their choice before making a terminal decision. Although the audio-visual presentation of this signal pattern appears to display the same degree of variability as mode 1, even though the SNR is higher, the calculated *P(C)* is the highest in this mode compared to all other treatments of amplitude modulated signal.

For signal pattern 1.2, the band of noise centered at 4000 Hz, the auditory presentation case showed the greatest degree of variability and also the highest SNR. It appears that in this signal presentation mode, subjects had difficulty in discriminating the signal. This is shown in the significantly higher response SNR for mode 1 over either mode 2 or 3 and also in mode 1 having the lowest calculated *P(C)* for this signal pattern in any presentation mode.

The same signal pattern in the visual-only presentation mode had a significantly lower SNR and significantly reduced variability. This would indicate that subjects were able to utilize the visual information very well in detecting and discriminating the signal pattern. This treatment of the signal pattern yielded the lowest SNR of the three presentation modes, although there was no significant difference between modes 2 and 3 in response SNR.

This signal pattern presented in the audio-visual mode again significantly reduced the variability compared with mode 2 and greatly reduced compared to mode 1. The SNR is slightly but not significantly higher than mode 2. The calculated *P(C)* in this case showed the subjects to be discriminating the signal pattern very well. The calculated *P(C)* in this case is unity and the 90% confidence interval is quite narrow. In the treatment of this signal pattern, the audio-visual presentation appears to yield the best performance for subjects in discriminating the signals.

For signal pattern 1.3, the band of noise centered at 1000 Hz, presented in the auditory mode only, the subjects' performance showed great variability in response SNR. In going to 2.3, the visual-only presentation case yielded a significant reduction in variability over mode 1, although the SNR at response is not significantly different. It would appear therefore that the subjects were able to detect the signal quite well in the auditory mode, but were able to detect and discriminate the signal more effectively in the visual mode; this is shown by reduction in variability. In the visual mode, the calculated *P(C)* was found to be unity, indicating a very high degree of discrimination performance for the subjects in this mode of presentation.

Signal pattern 3.3, the audio-visual presentation case, yields a further significant reduction in variability for the subjects' discrimination performance over the visual-only case; however, there was a slight reduction in the calculated *P(C)*.

It can be seen from the previous discussion and from the entries in Table 2 that the addition of the visual input of certain types of signals proves to be of benefit while for other types of signal patterns, it proves to be a hindrance. By comparing columns 4 and 6 of Table 2 for the different signal patterns across modes, we can quickly see which signal patterns yield the best performance in which presentation mode. For the case of amplitude modulation, signal patterns 1.1, 2.1, and 3.1, we can see that the combined audio-visual presentation yielded the best subject performance, in terms of *P(C)*. For signal patterns 1.2, 2.2, and 3.2, we find that again the combined audio-visual presentation is superior to either audio or visual presentation singly. For signal patterns 3.1, 3.2, and 3.3, we find the visual-only presentation mode superior to either auditory alone or combined audio-visual presentation.

Although the *P(C)*'s are not significantly different, we can see from column 5 of Table 2 that in the combined auditory-visual presentation mode, the response variability for all signals was reduced. This would indicate that for all the signals, the subjects were more consistent with their responses in the dual-mode case.

To examine these data in terms of the information processing model proposed in Chapter III, we can examine the signal pairs and presentation modes in terms of *P(C)* to assess the support of the model. By examining column 6 of Table 2 we can see which presentation mode elicits the best performance. By examining the *P(C)* we can see that the combined audio-visual presentation mode elicits the best performance and supports the model.

Signal pattern 2.1 shows the breakdown of the comparator stage, since in the visual-only presentation mode, amplitude modulation cannot be detected, the comparator stage can match the incoming information with either stored pattern. Therefore, there is a chance probability of obtaining the correct match. This chance probability is borne out by the *P(C)*.

Overall, it appears from the data that signal treatments 3.1, 3.2 and 3.3 yield the best performance and also best fit the proposed model. As stated earlier, the differences in neural conduction times for the auditory and visual modalities may impact the processing of bisensory information. The cueing theory model proposed by Nickerson (1973) claims that the auditory system alerts the visual system to the presence of incoming information, thereby raising the sensitivity of the system and aiding in detection. This aiding may follow the pattern of the statistical summation model proposed by Loveless, Brebner, and Hamilton (1970) which claims that the outputs of the two channels summate statistically to yield a higher probability of detection.

It seems reasonable, that the two above mentioned models may either operate to some degree together or may combine to form a hybrid system. The differences in neural conduction time of the auditory and visual systems would lend acceptance of the Nickerson (1973) cueing theory model. However, if it is assumed that the processing time for each modality is equal, or approximately equal, then the more rapidly arriving auditory information in essence would have to wait at the input of the comparator stage for the arrival of the slower visual information. Based on the findings of previous studies by Colquhoun (1975), Bruckner and Mc Grath (1961), and others that dual mode detection is superior to either unimodal detection, but less than the arithmetic sum of the single detections, it would seem that the proportionality model of Corcoran and Weening (1969) would apply. It is possible that the detections of the single modalities add proportionally to yield an increased detection in the bimodal presentation case. Since the auditory system would cue the visual, it would appear that the auditory system would carry the greatest weight in this proportion. From this model it is also clear that in the dual mode, more detections would be reported due to this cueing of one system by the other.

This cooperation of the two modalities would increase the likelihood of a detection in both systems and then pass the information on to the comparator. In the comparator stage, the signal pattern would be matched with the stored trace and a same/different classification made. The results of this same/different decision would then be passed on to the response stage and either a response is made or the signal sampling process is repeated.

It must be borne in mind that the results reported in this thesis are of discrimination tasks and not of vigilance or pure detection tasks. However, here it is necessary for the subjects to detect the presence of the dichotomous feature and to discriminate that signal from a previously learned signal. In that light detection is the proper term to use and detection literature is appropriate.

Areas where additional research efforts could be focused would be the use of multiple features. These features could be adjacent or nonadjacent bands of noise, with or without amplitude modulation. Martin (1978) investigated the use of multiple features and their interactions in an auditory only paradigm. However, it remains to be seen how these multiple interacting features would affect the detection and classification of signals in a bisensory presentation paradigm. Also additional research might focus on the use of an altered visual display system. The display used in this thesis consisted of a spectrum analyzer outputing the instantaneous spectra of the input signal. An interesting transform might be to use a spectrum analyzer outputing an exponential average of the input signal. This exponential average would reduce the masking effect of the background noise and should allow for enhanced detection and discrimination performance in the visual mode, which would also aid in the audio-visual performance.

CHAPTER VI

SUMMARY AND CONCLUSIONS

The main intent of the experiment reported in this thesis was to assess the utility of the use of totally redundant visual information in addition to auditory information in an attempt to aid in the detection and discrimination of noise-like sounds. Noise-like sounds are defined as sounds other than speech or music which can convey information to a listener. Detection and interpretation of such sounds play a considerable role in everyday life in alerting man to potential hazards in his environment.

Three pair of signals were used in this experiment. The signals were laboratory-generated, and the pairs differed in the presence or absence of a *dichotomous feature.* The dichotomous feature was defined as a signal pattern characteristic present in one member of the pair but absent in the other member. The dichotomous features used in this experiment consisted of an octave band of noise centered at 500 Hz and amplitude modulated by a 10 Hz square wave, an octave band of noise centered at 4000 Hz, and an octave band of noise centered at 1000 Hz. These signal pairs were tested in three experimental treatments; auditory presentation, visual presentation, and combined audio-visual presentation.

A review of relevant literature on bimodal presentation of auditory stimuli was provided in Chapter II. Also included in Chapter II is a brief discussion of the transforms made to yield a visual representation of the auditory signals and the different types of visual display systems used by various researchers. Chapter III presents a summary of some of the relevant information-processing models which have been postulated to apply to the dual sensory stimuli presentation case. Also detailed in Chapter III is postulated a feature extraction model which may be applied to signal pattern detection and discrimination. The assumptions of this model were:

1. That a dichotomous feature is present in the signal pattern,
2. That the characteristics of this feature, bandwidth and level, allow for reasonable detectability.

The model consists of a signal reception and encoding stage where the acoustic energy is received and transformed into neural impulses. The next stage of the model detects and isolates the dichotomous feature of the signal. The following stage compares the isolated dichotomous feature to a previously learned pattern stored in memory. Upon the basis of this comparison, the following stage makes a same/different decision and passes on the decision to the response stage which either terminates the problem or returns to the input stage for another sample of the acoustic energy, and the process is repeated, comparing the feature to a different stored pattern.

In Chapter IV, the experiment designed to test the utility of using bisensory input to an operator and the use of this model for predicting performance is detailed. The experiment conducted over a four-month period using five university students as subjects and three pairs of signals patterns as stimuli is also presented. The procedure used for the experiment is called the modified threshold technique and is also detailed in Chapter IV. There were two signals presented to the subjects in the procedure; these signals were denoted A and B, and one of them was presented superimposed in a white noise background. The signal-to-noise ratio was then increased slowly until the subject was able to detect and discriminate the signal pattern and was willing to make a terminal decision under the criterion of reasonably certain of the choice. These signals were presented aurally, visually, or combined in an audio-visual presentation mode. In each presentation mode, the measures of interest were the SNR at response and the probability of a correct response.

Certain portions of the generated data were deemed unreliable and as such were omitted from subsequent analysis. These data were data from training sessions, the first event of each group of six, and events where subjects failed to respond. The remaining data were pooled across subjects for each signal pair and treatment mode. Comparisons between the presentation modes lead to the following conclusions:

1. For the case where amplitude modulation is the dichotomous feature, the auditory sense has the greatest detection and

discrimination performance. A visual-only presentation of this signal pattern yields no discrimination information and yields performance at or near chance level. The presence of a visual display in addition to the auditory presentation of this type of signal serves as a distractor for the subject and hence degrades performance.

2. An octave band of noise centered at 4000 Hz as a dichotomous feature yields maximum detection and discrimination performance when presented both visually and aurally. In addition, detection of this signal pattern is superior for the visual-only mode compared to the auditory-only presentation mode, as shown by the lower mean SNR at response.

3. An octave band of noise centered at 1000 Hz, as a dichotomous feature yields maximum detection and discrimination performance in the combined audio-visual presentation mode by significantly reducing the response variability. The visual-only presentation mode is superior to the auditory mode also by significant reduction of the variance.

The results of these findings were also compared to the postulated information processing model. The data were seen to fit certain aspects of the model while, for other aspects, the model seemed inadequate. The signals used in this experiment by no means cover the endless number of possible signal types; in fact, these signals are a very small subset. Other areas of possible research have been suggested. These areas would include the use of different signal patterns with other dichotomous and multiple features and investigations of the stimulus presentation procedures and their effects on subject's performance.

REFERENCES

Adams, J. A. A closed-loop theory of motor training. *Journal of Motor Behavior,* 1971, **3**, 111-150.

Bernstein, I. E., Rose, R. and Ashe, V. Preparatory state effects in intersensory facilitation. *Psychomonic Science,* 1970, **19**, 113-114.

Brown, A. E. and Hopkins, H. K. Interaction of the auditory and visual sensory modalities. *Journal of the Acoustical Society of America,* 1966, **41**, 1-6.

Bruckner D. N. and Mc Grath, J. J. *A comparison of performance on single and dual mode vigilance tasks.* Technical Report No. 8. Human Factors Research Incorporated, Los Angeles, California. February 1961.

Colquhoun, W. P. Evaluation of auditory, visual and dual mode displays for prolonged sonar monitoring in repeated sessions. *Human Factors,* 1975, **17**, 425-437.

Corcoran, D. W. J. and Weening, D. L. On combination of evidence from eye and ear. *Ergonomics,* 1969, **12**, 383-394.

Cornell, P. T. Detection of an octave band of noise as a function of stimulus presentation, *The Pennsylvania State University, Masters Thesis,* March 1978.

Craig, A., Colquhoun, W. P., and Corcoran, D. W. J. Combining evidence presented simultaneously to the eye and ear: A comparison of some predictive models. *Perception and Psychophysics,* 1976, 19, 473-484.

Eriksen, C. W., and Hake, H. W. Multidimensional stimulus differences and accuracy of discrimination. *Journal of Experimental Psychology,* 1955, **50**, 153-160.

Freund, J. E. Mathematical Statistics, Englewood Cliffs, N. J. : Prentice-Hall, (1971)

Gaarder, K. A theoretical and experimental examination of auditory and visual interaction. *Perceptual and Motor Skills,* 1969, **29**, 23-33.

Green, D. M. Detection of multiple component signals in noise. *Journal of the Acoustical Society of America*, 1958, **30**, 904-911.

Green, D. M. Auditory detection of a noise signal. *Journal of the Acoustical Society of America*, 1960, **32**, 121-131.

Green, D. M. and Swets, J. A. Signal detection theory and psychophysics. New York: Wiely, 1966.

Halpern, J. *Task difficulty and channel redundancy in bisensory information processing.* Computer Image Corporation and the Department of Psychology of the University of Denver, July 1970.

Haywood, R. C., Leshowitz, B., Parkinson, S. R. and Eddowes, E. E., *Visual and auditory information process aspects of the acquisition of flying skill.* Arizona State University, Flying Training Division Air Force Human Resources Laboratory, Williams Air Force Base Arizona, December 1974. **(NTIS No. AD-A007721)**

Hodge, M. H. and Reid, L. S. The influence of similarity between relevant and irrelevant information upon a complex identification task. *Perception and Psychophysics*, 1971, **10**, 193-196.

Janota, C. P. An experimental treatment of auditory discrimination of complex noise-like sounds (Doctoral dissertation, The Pennsylvania State University, 1977). Dissertation Abstracts International, 1977, 8, 222-B. (University Microfilms No. 77-23,242)

Kohfeld, D. L. Effects of the intensity of auditory and visual ready signals on simple reaction time. *Journal of Experimental Psychology*, 1969, **82**, 88-95.

Loveless, N. E. *Signal detection with simultaneous visual and auditory presentation.* **Report No. 1027**, Air Ministry, Flying Personnel Research Committee, London England, 1957.

Loveless, N. E., Brebner, J. and Hamilton, P. Bisensory presentation of information. *Psychological Bulletin*, 1970, **73**, 161-195.

Martin, D. W. Discrimination of noise-like sounds involving multiple interacting features. *The Pennsylvania State University, Masters Thesis*, August 1978.

Nickerson, R. S. Intensity facilitation of reaction time: Energy summation or preparation enhancement. *Psychological Review*, 1973, **80**, 489-509.

Osborn, W. C., Sheldon, R. W. and Baker, R. A. Vigilance performance under conditions of redundant and nonredundant signal presentation. *Journal of Applied Psychology,* 1963, 47, 130-134.

Reed, S. K., Psychological processes in pattern recognition. New York: Academic Press, 1973.

Sage, G. H. Introduction to motor behavior: A neurophysiological approach. Reading: Addison-Wesley, 1977.

Schafer, T. H. and Shewmaker, G. A. *A comparative study of the audio, visual and audio-visual recognition differentials for pulses masked by random noise.* USN Electronics Laboratory, Report No. 372, 1953.

Stevens, S. S. Some similarities between hearing and seeing. *Laryngoscope,* 1958, 68, 508-527. (NTIS No. AD 215589)

Wells, F. D. *The investigation and simulation of dynamic interactive displays of acoustic signals.* Computer Image Corporation. Denver Colorado. July 1971. (NTIS No. AD-728056)

Woodhead, M. M., Carpenter, A. and Webster, J. C. *Perceiving sound by their visible spectra.* Royal Navy Personnel Research Committee, January 1973. (NTIS No. AD-A019-317)

APPENDIX A

INSTRUCTIONS TO SUBJECTS

Prior to participating in the experiment, each subject was given a description of the objectives of the experiment, the procedures to be used, and the rules regarding scheduling of test sessions. The subjects were shown how to operate the test apparatus and how to use the cassette recorder. The subjects were encouraged to ask questions at any time during the course of the experiment. In addition to the following specific instructions which appeared at the beginning of every audio tape, there were instructions which appeared on the response cassettes which were used for the first session following a mode change.

> The test sequence will consist of two signals presented without interfering noise. These signals will be denoted Signal A and Signal B. Signal A will be presented then Signal B. The signals will then be repeated in the sequence A then B. During the response period, which will be indicated by a green light on the response recorder, either Signal A or Signal B will be presented in a noise. The amount of noise will decrease slowly. The objective is to indicate your decision as to which signal you conclude is mixed with the noise. Indicate your choice by pressing the switch marked A if you decide that Signal A was mixed with the noise, or press the switch marked B if you decide that Signal B was mixed with the noise. You should indicate this decision as soon as you can under the condition that you are reasonably certain of your choice. For this series of experiments the tests are organized into groups of events. For all events of a group, the Signals A and B will be the same. The Signals A or B are presented randomly in the noise with each being equally likely. Now, please indicate your classification decision for the following cases.

Voice comments following the grouped events and leading into another group or terminating the test sessions were as follows.

This is the end of one group of tests. Another group of tests follows. For this group, the Signals A and B will be the same. However, these will generally not be the same signals as in the previous group. Please try to learn these signals without regard for other signals you may have heard during this experiment.

This is the end of another group of tests. In the group of events which follows, the Signals A and B are the same throughout. These signals will generally not be the same as those heard previously. Please try to learn these sounds independent of other signals you may have been exposed to in this experiment.

End of another group of events. The final group of events follows. For this group as in those before, the Signals A and B will be the same. The likelihood of Signal A being presented in the noise is the same as is the likelihood of Signal B being presented in the noise.
This concludes the test session. Thank you for your cooperation.

APPENDIX B

DOCUMENTS

AND BETWEEN-MODE INSTRUCTIONS

CONSENT FORM

____
Date

TITLE: An Investigation of Dual Sensory Presentation of Complex Noise-Like Sounds.

INVESTIGATOR: Alfred Barbour

PURPOSE: The purpose of this study is to test the ability of a subject to respond to changes in visual and auditory stimuli.

SUBJECT'S STATEMENT: The test will be conducted in a quiet enclosure using headphones and a visual display screen. The loudness of the signals presented over the headphones has been carefully controlled. At no time will the sound be so loud as to cause discomfort but, if at any time you should feel uncomfortable you should remove the headphones and leave the test enclosure. Your doing so will not be used as a basis for discontinuing your participation in the test. You may, however, terminate your participation in the experiment at any time.

For your convenience, test sessions may be scheduled at most reasonable times. However, to avoid fatigue and possible biasing of the experiment, no more than one 45 minute test session may be done in any 24 hour period.

The sounds which you will hear may or may not be familiar to you. The visual signals will be a representation of the sound. The method for recording your response, the order in which the signals are presented, and the details of the experiment, its overall objectives, its application and rationale will be explained to you prior to the beginning of any test session. If you have any questions, please ask for clarification by the experimenter.

I, __________________________, have read and understand this document.
subject signature

Witnessed by: ____________________ Date ________________

COMMENTS TO SUBJECTS PRIOR TO TESTING

In this experiment we are interested in the role of using more than one sensory modality to convey information to an operator. The experiment will be divided into three parts, designated Mode 1, Mode 2, and Mode 3. In Mode 1 we are examining the role of the auditory system. Therefore, you will be presented with auditory signals via headphones and be asked to make decisions concerning these signals.

In Mode 2 we are interested in the visual system and its role in processing information. In this mode you will receive a visual representation of the auditory signals, however you will not hear the signals. Again you will be asked to make decisions concerning these signals.

Mode 3 will a be combined effort, that is, you will receive both the auditory and visual signals. You will again be asked to make decisions concerning these signals.

The procedures to be followed in this experiment are as follows. The subject will activate a master power switch which will turn on all necessary equipment. The subject will then mount a test tape on the Crown 700 recorder. Following this the subject will load a designated cassette tape into the cassette player and activate the cassette in the record mode. The subject will then depress the "play" button on the Crown, enter the test booth, put on the headphones and the session will begin.

At the end of the test session the subject will remove the headphones, leave the test booth, rewind the test tape on the Crown, fast forward the cassette tape so that the flip side can be used, put away all tapes in their appropriate places and turn off the master power switch. This completes the test session.

During the changes from Mode 1 to Mode 2 and from Mode 2 to Mode 3, the first test session of each mode will have special instructions contained on the first cassette of the group. To hear these instructions the subject will proceed with the procedures described above, however upon loading the cassette into the player the subject will then depress the play button on the cassette player and listen to any special instructions. When the instructions are completed the subject will then rewind the cassette and proceed as usual. Special instructions will be required only during the first session following a mode change and there will be a note placed inside the appropriate cassette box attesting to the fact that the cassette contains special instuctions.

Due to the fact that mode changes require the experimenter to make changes in the experimental apparatus it is necessary that all subjects complete each mode before any one subject can move on to the next mode. Therefore it is important that all subjects try to complete each mode as quickly as possible and try not to skip or miss any days as this will inconvenience others by disprupting their schedules. Your cooperation and participation in this experiment is greatly appreciated.

Thank You.

Between-Mode Comments

Mode 2, Visual

During this portion of the experiment we are concerned with the role of the visual system in processing auditory information. For this part of the experiment it is not necessary for you to wear the headphones. Any voice comments or special instructions will be relayed to you via a speaker in the test enclosure. The signals of interest will be presented via the CRT display screen mounted outside the test enclosure and visible through the window. Other than the addition of the visual display, your task remains the same. That is, there will be two signals A and B presented visually without interferring noise. One of the signals will then be buried in the noise and the amount of noise will decrease slowly. You are to determine which signal was presented in the noise and respond appropriately. Your decision in this case however will be based on the visual information only since there will be no auditory information available. Now please rewind the cassette tape and begin the test session.

Mode 3, Audio/Visual

In this portion of the experiment we are concerned with the combined use of the auditory and visual systems in the processing of information. For this portion of the experiment you will once again be required to wear the headphones. The experimental procedure remains the same however. There will again be two signals and one will be presented in the noise. Only this time you will hear and see the signal. Your task is to use both the auditory and visual information to make a decision as to which signal is being presented and then make your response. Now please rewind the cassette tape and begin the test session.

AD-A074 936 PENNSYLVANIA STATE UNIV UNIVERSITY PARK APPLIED RESE--ETC F/G 5/10
AN INVESTIGATION OF DUAL SENSORY STIMULUS PRESENTATION OF COMPL--ETC(U)
JUN 79 A BARBOUR N00024-79-C-6043
UNCLASSIFIED ARL/PSU/TM-79-115 NL

2 OF 2
ADA 074936

END
DATE
FILMED
10 -79
DDC

OF 2

DA
074936

Co
Na
De
Wa

Co
Na
De
Wa

De
5C
Ca
Al

DISTRIBUTION

Commander (NSEA 09G32) Naval Sea Systems Command Department of the Navy Washington, DC 20362	Copies 1 and 2
Commander (NSEA 0342) Naval Sea Systems Command Department of the Navy Washington, DC 20362	Copies 3 and 4
Defense Documentation Center 5010 Duke Street Cameron Station Alexandria, VA 22314	Copies 5 through 16